SIMULATION AS A TOOL TO DEVELOP STATISTICAL UNDERSTANDING ®

Gail Burrill
Michigan State University
USA

Statistical reasoning is often presented through a variety of statistical "tests" – usually leaving many students bewildered. A foundation for understanding what statistical reasoning is and how it works can help students understand how to make sensible decisions from data before they move to formal techniques. Simulations, made possible by technology such as graphing calculators or computer software, can provide students with a conceptual basis for inference. By generating sampling distributions, students can analyze the behavior of a given statistic, explore whether a given observation is likely, investigate the effect of changing sample size, and consider how distributions differ. Such experiences give students a sense of how to reason from data and help to explain what is behind some of the formal tools of inference. Examples from the world outside of the classroom illustrate how simulation can be a tool in making sensible decisions giving students opportunities to see why statistics is important.

Business, industry, the government, the health field, education agencies collect data about what they do and how they operate to make decisions about what works and what does not, possible future directions, or how to become more efficient. Unfortunately, the answers are seldom clearly present in those data. Individuals, objects, repeated measurements on the same object vary, sometimes due to chance and sometimes due to other factors. Statistics is the science of reasoning from data in the presence of this variability (Moore, 1997). And a key element in this process is the role of randomization, or chance, as a tool in learning how to recognize, measure, decrease, or stabilize variability. Chance behavior, while unpredictable in the short run, has a regular and predictable pattern in the long run. Statistical inference, making use of this predictability, is based on an explicit chance model for the data (Freedman, Posani, & Purves, 1998).

In real life, you would not have the opportunity to repeat a study many times. Simulation, however, is a powerful tool that allows you to investigate what happens if the study could be repeated over and over again. Simulation allows students to gather information about what happens by chance and to use statistical reasoning to analyze the outcomes. Simulation can be used to create a chance model of the situation to explore the variability in the distribution of a sample statistic describing some characteristic of the population. Such sampling distributions answer the question: How would the statistic behave if the process were repeated many, many times? A key question is: does the observed behavior differ from what would be expected just by chance?

To understand how this reasoning process works and before setting up a theoretical hypothesis test, students need to experience for themselves how sampling distributions behave and to explore the patterns that emerge. Graphing calculators allow students to work on a daily basis in whatever place they choose to work to experiment with the data and test the conditions to see whether the patterns change. This paper presents three problems that can be analyzed by simulation. The objective is to create a process using simulation that will enable students to establish a sense of how to reason from data. More formal approaches to inferential statistics can be done later if and when they are appropriate for the course and the student. In each of the examples, the task is to define a statistic in response to the question of interest, generate a sampling distribution of the random behavior of this statistic; analyze the variability in the sampling distribution, and compare the observed behavior to the simulated random behavior. Probability is used as a tool to quantify how likely the observed behavior would occur if it were due to chance.

WESTVACO CASE

In the early 1990s, the Westvaco Corporation, due to changing circumstances, did not have enough work for all of its employees. Thus, they "laid off" some of the employees until such time, if ever, that the market becomes once again more favorable (Cobb, 1997). Westvaco laid off employees in five rounds, a different number of employees in different divisions each time. One of the employees who was let go, Robert Martin, became concerned that the company was laying off employees based on age, which would violate age-discrimination laws. As a consequence he sued the company in a court of law. In the third round of lay-offs, the ages of the three employees in one category who were laid off were 55, 55, and 64. To keep the problem manageable, suppose that the only other information you knew was the ages of all of the ten employees in this category: 25, 33, 35, 38, 48, 55, 55, 55, 56, 64. (In the actual case, other information such as seniority, work performance, and salary was also available.) Based just on this information, does it seem that Martin had a legitimate claim? In other words, were the ages of those laid off really older than those who were not laid off and, if so, could this have happened just by chance?

What could have occurred by chance? To find out, students can create a simulation model of the ten employees and use a chance mechanism to select those to lay off. One approach might be to choose without replacement from a set of ten random numbers each representing one of the ages. (In this case, note that if “1” represents an age of 25, “6”, “7”, and “8” would all represent the age 55.) A randomly chosen sample of three numbers will represent those who were laid off just by chance. How do these ages vary from the ages of those who were actually laid off? Students can choose a statistic that describes the sample of ages, then use simulation to create a sampling distribution for that statistic. In the sample below, the statistic was the mean age of the three laid off employees in a simulation of 116 rounds of layoffs.

Table 1. *Mean Age of Those Laid Off*

Mean age of those laid off	Frequency	Mean age of those laid off	Frequency
31		45	11111 11
32	1	46	11111 11111
33	11	47	11
34		48	11111 11111
35	11	49	1
36	11	50	11111 111
37	11	51	11111 11111
38	1111	52	11111 1
39	111	53	1111
40	1111	54	11
41	11111 11111	55	11111 1
42	11111 11	56	
43	111	57	11111 1
44		58	1111
		59	

The actual mean age of the three laid-off employees was 58. In the distribution of the mean ages from the samples, a mean age of 58 or more occurred 4 out of 116 or about 3% of the time. The decision – which separates statistics from mathematics – is whether you (or the judge) feel that something that occurs 3% of the time is an unlikely event. In most cases, an event that occurs less than 5% is “statistically significant.” It seems that a mean age of 58 does not often occur by chance; there seems to be reason to doubt that the lay-offs were due just to chance. (The actual case, which involved a more complete and thorough analysis of all of those laid off, was settled out of court.)

TRUE FALSE TESTS

What is the chance of passing a true-false test with a score of 60% just by guessing? Will your chance of passing increase, decrease, or stay the same if the number of questions increases?

Suppose a test had ten questions. The item of interest is the number correct in each test or set of ten questions. If 1 is correct and 0 incorrect, ten random numbers either 0 or 1 will represent each of the ten questions on the test where the number 1 is the number correct. Table 2 shows one sampling distribution of 50 simulations, and Figure 1 displays a histogram of the results.

Table 2. Ten-question true-false test

Number correct	Number of tests
0	
1	1
2	1111
3	11111
4	11111 11111 1
5	11111 1111
6	11111 1111
7	11111 11111
8	1
9	

According to this simulation, a score of 60% or better occurred 20 out of 50 times - a 40% chance of passing. It is important for students to see the distribution grow. They should think about what shape they expect to see and why. By inspecting the frequency distribution rather than just summary counts, students can have a visual image of what 40% probability looks like in relation to the distribution as well as experience thinking about probability as area.

Technology is a critical tool in enabling students to carry out the simulations but must be used carefully and only when students have done enough hands-on activities to appreciate what the technology generates. The commands (sum(randInt (0, 1, 10)) on a TI-83 will generate a set of ten random numbers either 0 or 1 and sum the set. Pressing ENTER will simulate the next test, i.e. generate the next set of ten numbers. The theoretical probability of passing at the 60% level could be found on a TI 83 by using 1-binomcdf (10, 5, 5)) or about 38% of the time, but just as in the earlier example, the formal work can come at a later time. Some students might be ready to use rand Bin(10, .5, 50)STO L1 and plot L1 to see the distribution. Care must be taken here to ensure that the goal of the lesson does not become which button to push and lose the meaning.

By exploring sampling distributions from simulations of tests with an increasing number of questions students can see what happens to the shape, mean, and standard deviation. The chance of getting a grade of 60% decreases as the number of questions increases (Figures 1, 2, Table 3).

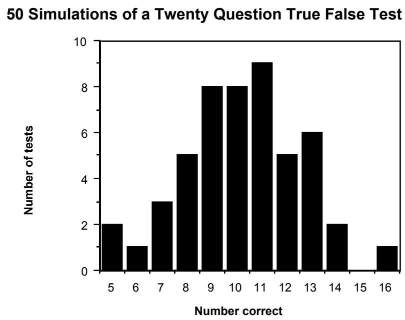

Figure 1 . Number correct in 20 questions

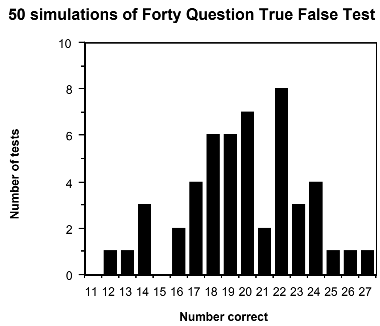

Figure 2. Number correct in 40 questions

Table 3. *Statistics for the Number of Correct Answers on True False Tests*

50 trials of	10 Questions	20 Questions	40 Questions
Mean no. correct	4.9	10.2	19.8
Standard deviation	1.69	2.35	3.36
Median	5	10	20
Interquartile range	2	3	4
Prob score >60%	40%	28%	14%

In all three simulations, the mean and the medians were near or at 50% correct. Students should notice, however, that as the number of questions increases, the distribution becomes more "stable". The variability measured by standard deviation and interquartile range decreased proportionally; for example, the interquartile ranges went from 1.7 out of 10 or 17% of the total range to 2.4 out of 20 or 12% to 3.4 out of 40 or 8.5% of the total range.

ANTIBIOTICS AND E-COLI

"Antibiotics can worsen E-coli complications."

According to a study from the University of Washington School of Medicine (Wong et al, 2000), children who may be infected with the bacteria E-coli 0157: H7 should not be treated with antibiotics because they raise the risk of a potentially deadly complication called hemolytic uremic syndrome (HUS). Researchers looked at 71 children with E-coli poisoning, nine of whom were treated with antibiotics. Of the nine, five developed HUS. Among the remaining 62, five developed HUS. Do the data support the headlines?

One approach to analyzing the problem is to consider the extreme cases. If there were strong evidence of a relationship between taking antibiotics and contacting HUS, all of those taking antibiotics would contact HUS (Table 4).

Table 4. *Strong evidence relating antibiotics and HUS*

	Antibiotics	No Antibiotics	Total
HUS	9	1	10
No HUS	0	61	61
Total	9	62	71

The other extreme would be to have none of those who received antibiotics contact HUS (Table 5).

Table 5. *No evidence relating antibiotics and HUS*

	Antibiotics	No Antibiotics	Total
HUS	0	10	10
No HUS	9	52	61
Total	9	62	71

If the relationship between taking antibiotics and contacting HUS was random, you would expect about the same effect for those taking antibiotics and for those not taking antibiotics. That is, the proportion of those taking antibiotics and contacting HUS would be the same as the proportion of those who did not take antibiotics yet contacted HUS (Table 6).

Table 6. *Expected relation between antibiotics and HUS due to chance*

	Antibiotics	No Antibiotics	Total
HUS	1.3	8.7	10
No HUS	7.7	53.3	61
Total	9	62	71

If all nine of those with E-coli who took antibiotics contacted HUS, some connection would seem to exist between taking antibiotics and contacting HUS. If none of the children who took antibiotics contacted HUS, there would be no reason to suspect a relationship. If the connection between antibiotics and HUS were random, you would expect one or two of the children who took antibiotics would have contacted HUDS (9 is to 71 as x is to 10). Students may have difficulty, however, when the evidence about taking antibiotics falls into a 'gray area'. Table 7 displays the original data.

Table 7. *E-Coli Study on the Relation of Antibiotics and HUS*

	Antibiotics	No Antibiotics	Total
HUS	5	5	10
No HUS	4	57	61
Total	9	62	71

Of the nine who took antibiotics, how likely is it that five of them would contact HUS when you expect one or two? In such cases, the relationship between antibiotics and HUS is not clearly obvious without further investigation. How would the number of E-Coli patients taking antibiotics who contacted HUS vary just by chance?

A bar graph (Figure 3) comparing those who contacted HUS and those who did not in the two categories, antibiotics and no antibiotics, can picture whether it is reasonable to suspect that a relation might exist.

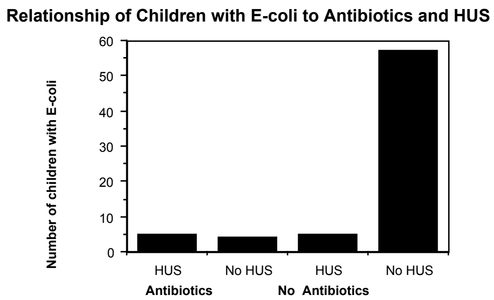

Figure 3.

The situation can be simulated by using two colors of marbles: ten black marbles to represent those with HUS and 61 red to represent those with no HUS. The structure of this simulation is based on the fact that 10 out of the 71 of those with E-coli contacted HUS while 61 of those with E-coli did not. That is, the row and column totals from the table are fixed. Count out 9 marbles for those who received antibiotics and count the number of black; count the number of black; for example, if three are black, then three of the children contacted HUS. The simulated counts generated by the chance process will produce a sampling distribution of counts for those who had antibiotics and contacted HUS. The observed number of patients for whom this happened (5) can then be compared to the sampling distribution produced by the simulation. In the sampling distribution of 50 simulations pictured in Figure 4, 5 did not occur at all.

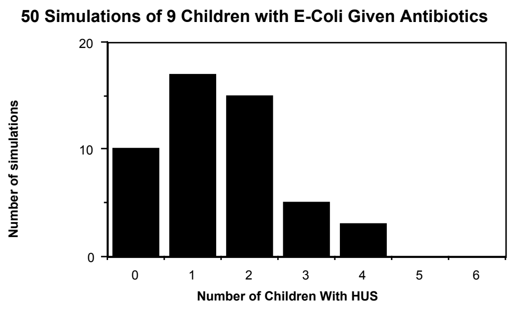

Figure 4.

The evidence seems to support the claim that contacting HUS was not due to chance. Because the observed number of those taking antibiotics who also contacted HUS was very unlikely to occur by chance, the data do seem to support the headlines.

Questions about the design of the study itself, however, might be raised and are important to consider before making any real claims about the validity of the conclusions.

CONCLUSION

Research indicates that students need to see that they are learning something useful and relevant and are motivated when they can use the information they learn to do something that will have an effect on others (McCombs, 1996; Pintrich & Schunk, 1996). Students learn when they are actively involved in choosing and evaluating strategies, considering assumptions, and receiving feedback. They often fail to connect everyday knowledge to subjects taught in school (NRC, 1999). The problems above give students these opportunities. The paper did not explicitly explore the mathematics necessary to carry out the work in analyzing the three problems, but it is considerable and includes manipulating formulas, working with symbols, creating models, making and interpreting graphs. Exploring problems such as these makes mathematics and statistics come alive. The work supports the research by illustrating that statistical reasoning can be a powerful tool in making sense of real problems.

REFERENCES

Cobb, G., & Cryer, J. (1997). *An electronic companion to statistics*. San Francisco, CA.: Cogito Learning Media, Inc.

Freedman, D., Posani, R., & Purves, R. (1998). *Statistics*. Third Ed. (pp. 161-170). New York, NY: W. W. Norton & Company, Inc.

McCombs, B.L. (1996). Alternative perspectives for motivation. *Developing engaged readers in school and home communities.* Baker, L., Afflerluck, P., & Reinking, D. (Eds.) Mahwah NJ: Erlbaum. (Pp. 67-87).

Moore, D. (1997). *Statistics: Concepts and controversies*. New York, NY: W.H. Freeman and Company.

National Research Council (2000). *How people learn: Brain, mind, experience, and school*. Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) Committee on Developments in the Science of Learning, National Research Council. (p. 57). Washington DC: National Academy Press.

Wong, C. S., Jelacic, S., Habeeb, R. L., Watkins, S. L., & Tarr, P. I. (2000). The risk of the hemolytic-uremic syndrome after antibiotic treatment of escherichia coli O157:H7 infections. *The New England Journal of Medicine 342*, 1930-6.